MIST

MOLECULAR INFORMATION STORAGE

INTELLIGENCE VALUE

The MIST program seeks to use sequence-controlled polymers (e.g., DNA) as the basis for deployable digital information storage technologies. These technologies can eventually scale into the exabyte regime and beyond with reduced physical footprint, power, and cost requirements relative to conventional storage technologies. If successful, this program would empower the Intelligence Community to meet its future data storage needs without facing prohibitive resource constraints.

The scale and complexity of the world's data problems are increasing rapidly. Use cases that require storage and random access from exabytes (10^{18} bytes) of mostly unstructured data are now well-established in the private sector and are of increasing relevance to the public sector. However, meeting these requirements poses extraordinary logistical and financial challenges. For example, today's exabyte-scale data centers occupy large warehouses, consume megawatts of power, and cost hundreds of millions of dollars to build, operate, and maintain over their lifetimes. This resource-intensive model does not offer a tractable path to scaling beyond the exabyte regime in the future.

The MIST program launched in September 2019 to deliver a new category of resource-efficient storage technologies to the IC. At the conclusion of this 4-year program, MIST systems are aiming to demonstrate the ability to write 1 Terabyte, and read 10 Terabytes with random access capabilities, per day at a cost of less than $1,000 while consuming less than 1 kilowatt of power, and fitting on a table-top. MIST seeks to accomplish this by using sequence-controlled polymers (e.g., controlled DNA synthesis) as a data storage medium, and by building the necessary devices and information systems to interface with this medium.
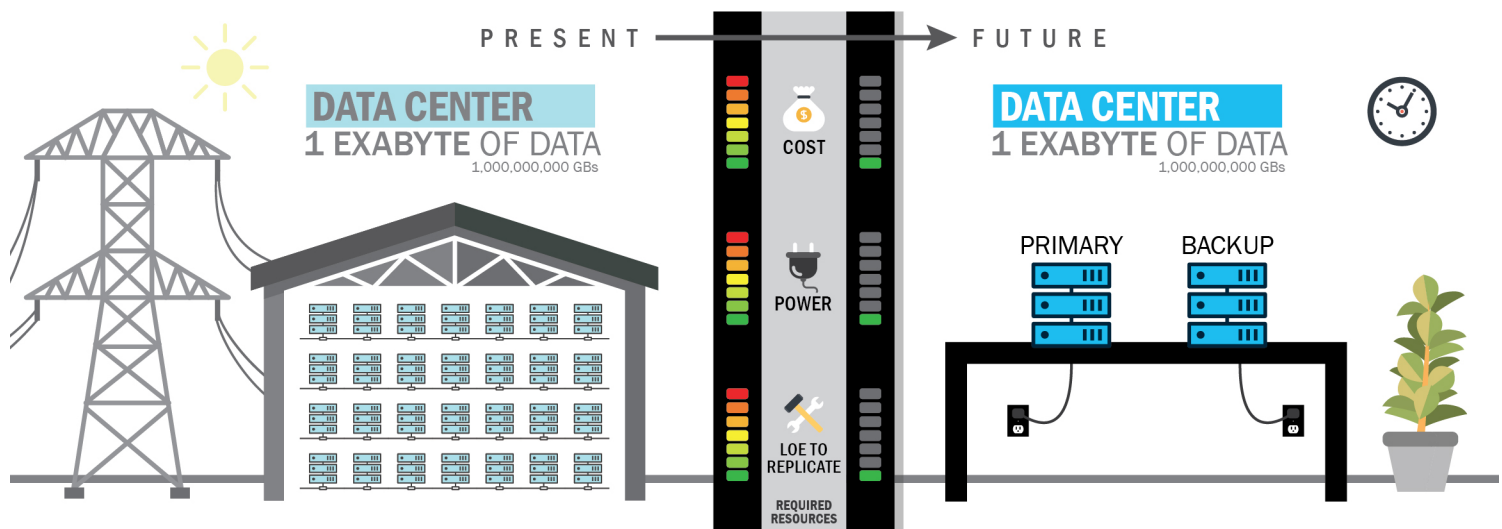
PRIME PERFORMERS

- Georgia Tech Research Institute
- Broad Institute

TESTING AND EVALUATION PARTNERS

- Los Alamos National Laboratories
- Sandia National Laboratories
- U.S. Army Futures Command
- Army Research Laboratory-South

KEYWORDS

- Data storage
- DNA synthesis and sequencing

PRESENT → FUTURE

DATA CENTER
1 EXABYTE OF DATA
1,000,000,000 GBs

COST

POWER

LOE TO REPLICATE

REQUIRED RESOURCES

DATA CENTER
1 EXABYTE OF DATA
1,000,000,000 GBs

PRIMARY   BACKUP

The MIST program seeks to reduce the resource requirements of today's data centers by several orders of magnitude. In the future, DNA media and associated write/read devices will make it possible to maintain, copy, and transport exabyte-scale data archives in a table-top form factor

MIST

PROGRAM MANAGER
Pedro Espina Ph.D.
Phone: 301-243-2089
pedro.espina@iarpa.gov

www.iarpa.gov

@IARPAnews

Latest Revision: 05/17/2023